



IFWO

RAW SEQUENCE LISTING

DATE: 08/30/2004

PATENT APPLICATION: US/10/780,731

TIME: 13:09:30

Input Set : A:\10-780,731 Sequence Listing.txt

Output Set: N:\CRF4\08302004\J780731.raw

```

3 <110> APPLICANT: National Institute of Advanced Industrial Science and
4     Technology
5     ISHIKAWA, Kazuhiko
6     JOEN, Sung-Jong
8 <120> TITLE OF INVENTION: HEAT-RESISTANT DNA LIGASE
10 <130> FILE REFERENCE: Q79424
12 <140> CURRENT APPLICATION NUMBER: 10/780,731
13 <141> CURRENT FILING DATE: 2004-02-19
15 <150> PRIOR APPLICATION NUMBER: JP 2003-045224
16 <151> PRIOR FILING DATE: 2003-02-24
18 <160> NUMBER OF SEQ ID NOS: 7
20 <170> SOFTWARE: PatentIn version 3.3
22 <210> SEQ ID NO: 1
23 <211> LENGTH: 1860
24 <212> TYPE: DNA
25 <213> ORGANISM: Aeropyrum pernix
28 <220> FEATURE:
29 <221> NAME/KEY: CDS
30 <222> LOCATION: (1)..(1860)
32 <400> SEQUENCE: 1
33 gtg ggg tgt ctg gtt ttg gct tct agc tct ggg ggt gtt gga ggc ggt      48
34 Val Gly Cys Leu Val Leu Ala Ser Ser Ser Gly Gly Val Gly Gly Gly
35 1          5          10          15
37 gac atg cct ttc aag ccc gtg gct gag gcc ttc gcc tcc atg gag agg      96
38 Asp Met Pro Phe Lys Pro Val Ala Glu Ala Phe Ala Ser Met Glu Arg
39          20          25          30
41 ata acc tct agg acg cag ctc acc ctc ctc ctc aca agg ctc ttc aag      144
42 Ile Thr Ser Arg Thr Gln Leu Thr Leu Leu Leu Thr Arg Leu Phe Lys
43          35          40          45
45 tcc acg ccc ccg ggg gcg atc ggt att gtg gtg tac ctg atc cag ggg      192
46 Ser Thr Pro Pro Gly Ala Ile Gly Ile Val Val Tyr Leu Ile Gln Gly
47          50          55          60
49 aag ctg ggg ccc gac tgg aag ggg ctg ccg gag ctg ggt gtc ggg gag      240
50 Lys Leu Gly Pro Asp Trp Lys Gly Leu Pro Glu Leu Gly Val Gly Glu
51 65          70          75          80
53 aag ctg ctt gta aag gcc ata gcc ctg gct tac aag gcc act gag gag      288
54 Lys Leu Leu Val Lys Ala Ile Ala Leu Ala Tyr Lys Ala Thr Glu Glu
55          85          90          95
57 agg gtt gag agg ctc tac aag tct gta ggc gac ctg ggg agt gtg gcc      336
58 Arg Val Glu Arg Leu Tyr Lys Ser Val Gly Asp Leu Gly Ser Val Ala
59          100         105         110
61 gag agg ctg tcg cgg gag tac cgc tcc cgg gct gcc agg gcc gtc acc      384
62 Glu Arg Leu Ser Arg Glu Tyr Arg Ser Arg Ala Ala Arg Ala Val Thr

```

ENTERED

RAW SEQUENCE LISTING

DATE: 08/30/2004

PATENT APPLICATION: US/10/780,731

TIME: 13:09:30

Input Set : A:\10-780,731 Sequence Listing.txt

Output Set: N:\CRF4\08302004\J780731.raw

63	115	120	125	
65	ctg gag gcg ttc atg gcg gga ggg ggg gag gcg ctg act gtg agg agg	432		
66	Leu Glu Ala Phe Met Ala Gly Gly Gly Glu Ala Leu Thr Val Arg Arg			
67	130	135	140	
69	ggt tac aac acg ctg tac agg ata gcc atg gcg cag ggt gag ggg agc	480		
70	Val Tyr Asn Thr Leu Tyr Arg Ile Ala Met Ala Gln Gly Glu Gly Ser			
71	145	150	155	160
73	agg gac atc aag ctt agg ctg ctg gcc ggc ctc ctg gcg gac gcc gag	528		
74	Arg Asp Ile Lys Leu Arg Leu Leu Ala Gly Leu Leu Ala Asp Ala Glu			
75	165	170	175	
77	ccc gtg gag gcg aag tat att gtg agg ttt gtg gag ggg agg ctg agg	576		
78	Pro Val Glu Ala Lys Tyr Ile Val Arg Phe Val Glu Gly Arg Leu Arg			
79	180	185	190	
81	gtg ggt gtt ggg gac gcg acc gtc ctc gac gcc ctc gcc atg gcc ttc	624		
82	Val Gly Val Gly Asp Ala Thr Val Leu Asp Ala Leu Ala Met Ala Phe			
83	195	200	205	
85	ggc ggc ggg gcc cac gcg agg ccc gtt ata gag agg gcc tac aac ctc	672		
86	Gly Gly Gly Ala His Ala Arg Pro Val Ile Glu Arg Ala Tyr Asn Leu			
87	210	215	220	
89	agg gcc gac cta ggc tac ata gcg gag gtc gtg gcc agg gag ggt gtt	720		
90	Arg Ala Asp Leu Gly Tyr Ile Ala Glu Val Val Ala Arg Glu Gly Val			
91	225	230	235	240
93	gat gcg ctg agg ggt gtg aag ccc cag gtc ggc gtt cct ata agg ccg	768		
94	Asp Ala Leu Arg Gly Val Lys Pro Gln Val Gly Val Pro Ile Arg Pro			
95	245	250	255	
97	atg ctg gcc gag agg ggg agg gac ccg gct gag ata ctc agg aag gtg	816		
98	Met Leu Ala Glu Arg Gly Arg Asp Pro Ala Glu Ile Leu Arg Lys Val			
99	260	265	270	
101	ggg ggc agg gct gtc gtc gag tat aag tac gat ggg gag agg gcg cag	864		
102	Gly Gly Arg Ala Val Val Glu Tyr Lys Tyr Asp Gly Glu Arg Ala Gln			
103	275	280	285	
105	ata cac aag aag gac ggg gag gtc tac atc tac tcg agg agg ctt gag	912		
106	Ile His Lys Lys Asp Gly Glu Val Tyr Ile Tyr Ser Arg Arg Leu Glu			
107	290	295	300	
109	aac ata acc agg atg ttc ccc gac gtg gtt gag atg gcg agg aag ggc	960		
110	Asn Ile Thr Arg Met Phe Pro Asp Val Val Glu Met Ala Arg Lys Gly			
111	305	310	315	320
113	ctc aaa gcc ggg gag gct ata gtc gag ggg gag ata gtg gcc gta gac	1008		
114	Leu Lys Ala Gly Glu Ala Ile Val Glu Gly Glu Ile Val Ala Val Asp			
115	325	330	335	
117	cca gac aac tat gag ata cag ccc ttc cag gtc ctc atg cag agg aag	1056		
118	Pro Asp Asn Tyr Glu Ile Gln Pro Phe Gln Val Leu Met Gln Arg Lys			
119	340	345	350	
121	agg aag cac gac ata cac agg gtc atg agg gag gtg ccc gtc gcc gtc	1104		
122	Arg Lys His Asp Ile His Arg Val Met Arg Glu Val Pro Val Ala Val			
123	355	360	365	
125	ttc ctc ttc gac gcc ctc tac gtg gac ggc gag gac ctc aca agc aaa	1152		
126	Phe Leu Phe Asp Ala Leu Tyr Val Asp Gly Glu Asp Leu Thr Ser Lys			
127	370	375	380	

RAW SEQUENCE LISTING

DATE: 08/30/2004

PATENT APPLICATION: US/10/780,731

TIME: 13:09:30

Input Set : A:\10-780,731 Sequence Listing.txt

Output Set: N:\CRF4\08302004\J780731.raw

```

129 ccc ctc ccc gag agg cgc agg agg ctc aag gag ata gtt gtg gag acg      1200
130 Pro Leu Pro Glu Arg Arg Arg Arg Leu Lys Glu Ile Val Val Glu Thr
131 385                               390                               395                               400
133 ccc ctc tgg agg ctg gcg gag tcc atc gag acc agc gac ccc gag gag      1248
134 Pro Leu Trp Arg Leu Ala Glu Ser Ile Glu Thr Ser Asp Pro Glu Glu
135                               405                               410                               415
137 ctg tgg acc ttc ttc ctg aag gcc ata gag gag ggg gcc gag ggc gtc      1296
138 Leu Trp Thr Phe Phe Leu Lys Ala Ile Glu Glu Gly Ala Glu Gly Val
139                               420                               425                               430
141 atg gtc aag gcc gtc cac agg gac tca gtc tac acc gcg ggc gta agg      1344
142 Met Val Lys Ala Val His Arg Asp Ser Val Tyr Thr Ala Gly Val Arg
143                               435                               440                               445
145 ggg tgg ctg tgg gtc aag ctg aag agg gat tac aag agc gag atg atg      1392
146 Gly Trp Leu Trp Val Lys Leu Lys Arg Asp Tyr Lys Ser Glu Met Met
147                               450                               455                               460
149 gac acg gtg gac ctc gta gtg gtg ggc gcc ttc tac ggc agg ggg aag      1440
150 Asp Thr Val Asp Leu Val Val Val Gly Ala Phe Tyr Gly Arg Gly Lys
151 465                               470                               475                               480
153 agg ggc ggg aag ctc agc agc ctg ctc atg gcc gcc tac gac cca gac      1488
154 Arg Gly Gly Lys Leu Ser Ser Leu Leu Met Ala Ala Tyr Asp Pro Asp
155                               485                               490                               495
157 agg gac gtg ttc ccc acc gtc tgc aag gtg gcc aca ggg ttc acg gac      1536
158 Arg Asp Val Phe Pro Thr Val Cys Lys Val Ala Thr Gly Phe Thr Asp
159                               500                               505                               510
161 gag gag ctg gac agg atg aac gag atg ctg aag aag cac atc ata ccc      1584
162 Glu Glu Leu Asp Arg Met Asn Glu Met Leu Lys Lys His Ile Ile Pro
163                               515                               520                               525
165 agg aag cac ccg agg gta gag tgc agg ata gag cct gac gtg tgg gtg      1632
166 Arg Lys His Pro Arg Val Glu Ser Arg Ile Glu Pro Asp Val Trp Val
167                               530                               535                               540
169 gag ccc gcc ctc gtg gcg gag ata ctg ggc gcc gag ctc acc ctc tca      1680
170 Glu Pro Ala Leu Val Ala Glu Ile Leu Gly Ala Glu Leu Thr Leu Ser
171 545                               550                               555                               560
173 cca atg cac acc tgc tgc ctc aac act gtg agg ccg ggg gtg ggg ata      1728
174 Pro Met His Thr Cys Cys Leu Asn Thr Val Arg Pro Gly Val Gly Ile
175                               565                               570                               575
177 agc ata agg ttc ccc agg ttc ata agg tgg agg gac gac aag agt ccg      1776
178 Ser Ile Arg Phe Pro Arg Phe Ile Arg Trp Arg Asp Asp Lys Ser Pro
179                               580                               585                               590
181 gag gac gcg aca aca acc cac gag ctg ctc gag atg tac aag agg cag      1824
182 Glu Asp Ala Thr Thr Thr His Glu Leu Leu Glu Met Tyr Lys Arg Gln
183                               595                               600                               605
185 ttg agg agg gtt gaa gag ccg gcg gag cag gtg tag      1860
186 Leu Arg Arg Val Glu Glu Pro Ala Glu Gln Val
187                               610                               615
190 <210> SEQ ID NO: 2
191 <211> LENGTH: 619
192 <212> TYPE: PRT
193 <213> ORGANISM: Aeropyrum pernix

```

RAW SEQUENCE LISTING

DATE: 08/30/2004

PATENT APPLICATION: US/10/780,731

TIME: 13:09:30

Input Set : A:\10-780,731 Sequence Listing.txt

Output Set: N:\CRF4\08302004\J780731.raw

195 <400> SEQUENCE: 2

```

197 Val Gly Cys Leu Val Leu Ala Ser Ser Ser Gly Gly Val Gly Gly Gly
198 1          5          10          15
201 Asp Met Pro Phe Lys Pro Val Ala Glu Ala Phe Ala Ser Met Glu Arg
202          20          25          30
205 Ile Thr Ser Arg Thr Gln Leu Thr Leu Leu Thr Arg Leu Phe Lys
206          35          40          45
209 Ser Thr Pro Pro Gly Ala Ile Gly Ile Val Val Tyr Leu Ile Gln Gly
210          50          55          60
213 Lys Leu Gly Pro Asp Trp Lys Gly Leu Pro Glu Leu Gly Val Gly Glu
214 65          70          75          80
217 Lys Leu Leu Val Lys Ala Ile Ala Leu Ala Tyr Lys Ala Thr Glu Glu
218          85          90          95
221 Arg Val Glu Arg Leu Tyr Lys Ser Val Gly Asp Leu Gly Ser Val Ala
222          100          105          110
225 Glu Arg Leu Ser Arg Glu Tyr Arg Ser Arg Ala Ala Arg Ala Val Thr
226          115          120          125
229 Leu Glu Ala Phe Met Ala Gly Gly Gly Glu Ala Leu Thr Val Arg Arg
230          130          135          140
233 Val Tyr Asn Thr Leu Tyr Arg Ile Ala Met Ala Gln Gly Glu Gly Ser
234 145          150          155          160
237 Arg Asp Ile Lys Leu Arg Leu Leu Ala Gly Leu Leu Ala Asp Ala Glu
238          165          170          175
241 Pro Val Glu Ala Lys Tyr Ile Val Arg Phe Val Glu Gly Arg Leu Arg
242          180          185          190
245 Val Gly Val Gly Asp Ala Thr Val Leu Asp Ala Leu Ala Met Ala Phe
246          195          200          205
249 Gly Gly Gly Ala His Ala Arg Pro Val Ile Glu Arg Ala Tyr Asn Leu
250          210          215          220
253 Arg Ala Asp Leu Gly Tyr Ile Ala Glu Val Val Ala Arg Glu Gly Val
254 225          230          235          240
257 Asp Ala Leu Arg Gly Val Lys Pro Gln Val Gly Val Pro Ile Arg Pro
258          245          250          255
261 Met Leu Ala Glu Arg Gly Arg Asp Pro Ala Glu Ile Leu Arg Lys Val
262          260          265          270
265 Gly Gly Arg Ala Val Val Glu Tyr Lys Tyr Asp Gly Glu Arg Ala Gln
266          275          280          285
269 Ile His Lys Lys Asp Gly Glu Val Tyr Ile Tyr Ser Arg Arg Leu Glu
270          290          295          300
273 Asn Ile Thr Arg Met Phe Pro Asp Val Val Glu Met Ala Arg Lys Gly
274 305          310          315          320
277 Leu Lys Ala Gly Glu Ala Ile Val Glu Gly Glu Ile Val Ala Val Asp
278          325          330          335
281 Pro Asp Asn Tyr Glu Ile Gln Pro Phe Gln Val Leu Met Gln Arg Lys
282          340          345          350
285 Arg Lys His Asp Ile His Arg Val Met Arg Glu Val Pro Val Ala Val
286          355          360          365
289 Phe Leu Phe Asp Ala Leu Tyr Val Asp Gly Glu Asp Leu Thr Ser Lys
290          370          375          380

```

RAW SEQUENCE LISTING

DATE: 08/30/2004

PATENT APPLICATION: US/10/780,731

TIME: 13:09:30

Input Set : A:\10-780,731 Sequence Listing.txt

Output Set: N:\CRF4\08302004\J780731.raw

```

293 Pro Leu Pro Glu Arg Arg Arg Arg Leu Lys Glu Ile Val Val Glu Thr
294 385                      390                      395                      400
297 Pro Leu Trp Arg Leu Ala Glu Ser Ile Glu Thr Ser Asp Pro Glu Glu
298                      405                      410                      415
301 Leu Trp Thr Phe Phe Leu Lys Ala Ile Glu Glu Gly Ala Glu Gly Val
302                      420                      425                      430
305 Met Val Lys Ala Val His Arg Asp Ser Val Tyr Thr Ala Gly Val Arg
306                      435                      440                      445
309 Gly Trp Leu Trp Val Lys Leu Lys Arg Asp Tyr Lys Ser Glu Met Met
310                      450                      455                      460
313 Asp Thr Val Asp Leu Val Val Val Gly Ala Phe Tyr Gly Arg Gly Lys
314 465                      470                      475                      480
317 Arg Gly Gly Lys Leu Ser Ser Leu Leu Met Ala Ala Tyr Asp Pro Asp
318                      485                      490                      495
321 Arg Asp Val Phe Pro Thr Val Cys Lys Val Ala Thr Gly Phe Thr Asp
322                      500                      505                      510
325 Glu Glu Leu Asp Arg Met Asn Glu Met Leu Lys Lys His Ile Ile Pro
326                      515                      520                      525
329 Arg Lys His Pro Arg Val Glu Ser Arg Ile Glu Pro Asp Val Trp Val
330                      530                      535                      540
333 Glu Pro Ala Leu Val Ala Glu Ile Leu Gly Ala Glu Leu Thr Leu Ser
334 545                      550                      555                      560
337 Pro Met His Thr Cys Cys Leu Asn Thr Val Arg Pro Gly Val Gly Ile
338                      565                      570                      575
341 Ser Ile Arg Phe Pro Arg Phe Ile Arg Trp Arg Asp Asp Lys Ser Pro
342                      580                      585                      590
345 Glu Asp Ala Thr Thr Thr His Glu Leu Leu Glu Met Tyr Lys Arg Gln
346                      595                      600                      605
349 Leu Arg Arg Val Glu Glu Pro Ala Glu Gln Val
350                      610                      615
353 <210> SEQ ID NO: 3
354 <211> LENGTH: 21
355 <212> TYPE: DNA
356 <213> ORGANISM: Artificial Sequence
358 <220> FEATURE:
359 <223> OTHER INFORMATION: PCR primer
361 <400> SEQUENCE: 3
362 ggctgtctgg ttttggttc t
365 <210> SEQ ID NO: 4
366 <211> LENGTH: 27
367 <212> TYPE: DNA
368 <213> ORGANISM: Artificial Sequence
370 <220> FEATURE:
371 <223> OTHER INFORMATION: PCR primer
373 <400> SEQUENCE: 4
374 gtgaaggat cttacacct gctccgc
377 <210> SEQ ID NO: 5
378 <211> LENGTH: 35
379 <212> TYPE: DNA

```

21

27

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/780,731

DATE: 08/30/2004

TIME: 13:09:31

Input Set : A:\10-780,731 Sequence Listing.txt

Output Set: N:\CRF4\08302004\J780731.raw